

Data Collection and Preprocessing Phase

Date	12 June 2025
Team ID	SWTID1749627644
Project Title	Human Resource Management: Predicting Employee Promotions using Machine Learning
Maximum Marks	2 Marks

Data Quality Report:

The Data Quality Report Template will summarize data quality issues from the selected source, including severity levels and resolution plans. It will aid in systematically identifying and rectifying data discrepancies.

Data Source	Data Quality Issue	Severity	Resolution Plan
Skill Wallet Dataset	<ol style="list-style-type: none"> Missing values in the 'education', 'previous_year_rating' columns. Outliers in the 'length_of_service' column. Imbalanced classes in the 'is_promoted' column. 	Moderate	<ol style="list-style-type: none"> Fill missing values with the mode for education and previous_year_rating. Cap outliers in length_of_service using the upper bound (IQR method). Balance is_promoted classes using SMOTE oversampling.